# STOCK PRICE PREDICTIONS

# Hello!

Done By: **Zeinelabdin Salih**

Instructor: **Vijay Kumar**

March 27, 2021

| Page | Content | Slides |
|------|---------|--------|
| 1 | Title / Name / Date | 1 - 2 |
| 2 | Outline | 3 |
| 3 | Introduction / Executive Summary | 4 - 5 |
| 4 | Project Obejtives | 6 |
| 5 | Hypotheses Generation | 7 - 8 |
| 6 | Methodology | 9 – 10 |
| 7 | Descriptive Analytics | 11 - 16 |
| 8 | Assumptions | 17 - 18 |
| 9 | Predication Models | 19 - 23 |
| 11 | Conclusions | 24 |

# Introduction

## Dataset Brief:

- This dataset is mainly created for analyzing the historical stock price of TSLA. Starting from 2010 until February 2021.

- It contains 7 Variables and 2679 datapoints.

- The features of analysis are Date, High, Low, Open, Close, Volume and Adj Close

- It can be used to make cool charts of the stock price for learning data analysis for finance.

- Our Target variable is "**Close**" Price

- Data source is Kaggle Website

# Executive Summary

- I checked missing values and possibility of outliers In exploration analysis

- Since I am dealing with Time Series data, there are some assumption you have taking caring of it

- I used the statistical methods and some graph to explain those assumptions

- I used Arima (Auto Regressive Moving Average), Decision Tree and Linear Regression Models to predict the forecast for Tesla stock close price

# Project Objectives

- Analysis and find the relationship between variables
- Build models to predict Tesla Stock Close Price
- Observe which model is best for predication
- Compare forecast Price with exiting Stock price on Tesla website

# Hypothesis Generation

## Stock Price Hypothesis Attributes:

- **Data & Time:** which is the much important feature for stocks

- **Volume of Stocks:** Explaining the demand of stocks on daily bases

- **Open Price:** Considered as one of major factors, especially at beginning of the week

- **High & Low Price:** The highest and lowest price stated during a day

- **Close Price:** The super significance attribute, which is used for finding historical insights in order to predict the future forecasting

- **Adjusted Close Price:** Important because it shows the stock's value after dividends are posted

- **Other Factors :** The stock market and price is very sensitive to many aspects for Exp. Profitability of company, Inflation, Investors Movement, Natural Disasters and Pandemics.

# Hypothesis Generation

## Possible Outcomes

1. Do the close prices in coming two months it will raise up again?
2. Will stock volume decrease and drop down the close price as well ?

# Methodology

- The dataset available as structured data in CSV format
- I changed the "Date" variable type to date
- There no missing values detected
- The boxplots graphs showing there some outliers, but In my opinion, I don`t believe that we must treat outliers since we need to check the data stationary
- Project done in **Jupytor Notebook –**

  **Python 3.8**

```
Date          0
High          0
Low           0
Open          0
Close         0
Volume        0
Adj Close     0
dtype: int64
```

| | Date | High | Low | Open | Close | Volume | Adj Close |
|---|---|---|---|---|---|---|---|
| count | 2679 | 2679.000000 | 2679.000000 | 2679.000000 | 2679.000000 | 2.679000e+03 | 2679.000000 |
| unique | 2679 | NaN | NaN | NaN | NaN | NaN | NaN |
| top | 2013-06-04 | NaN | NaN | NaN | NaN | NaN | NaN |
| freq | 1 | NaN | NaN | NaN | NaN | NaN | NaN |
| mean | NaN | 71.456871 | 68.343933 | 69.925943 | 70.028605 | 3.176344e+07 | 70.028605 |
| std | NaN | 127.593072 | 121.612343 | 124.658055 | 125.036783 | 2.914939e+07 | 125.036783 |
| min | NaN | 3.326000 | 2.996000 | 3.228000 | 3.160000 | 5.925000e+05 | 3.160000 |
| 25% | NaN | 7.691000 | 7.414000 | 7.581000 | 7.575000 | 1.130900e+07 | 7.575000 |
| 50% | NaN | 45.495998 | 44.000000 | 44.650002 | 44.646000 | 2.503650e+07 | 44.646000 |
| 75% | NaN | 62.341000 | 60.201000 | 61.339001 | 61.503000 | 4.107750e+07 | 61.503000 |
| max | NaN | 900.400024 | 871.599976 | 891.380005 | 883.090027 | 3.046940e+08 | 883.090027 |

# Methodology

## Statistical Methods / Techniques Used:

- Installed new environut libraires

- Univariate, Multivariate Analysis

- User define, Lambda functions

- Assumptions Checking: Autocorrelation, Stationary

- Predictive Analysis: Arima, Linear Regression and Decision Tree Models
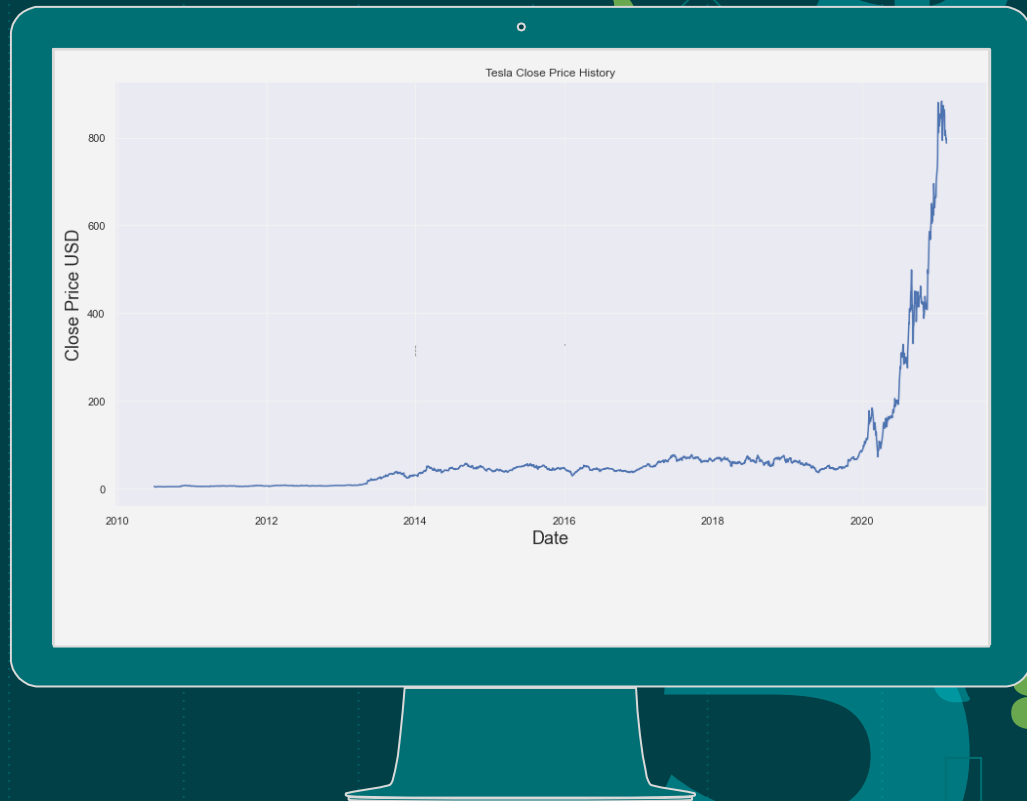
# Descriptive Analysis

─○ **Univariate Analysis**

Showing Outliers in Volume varible
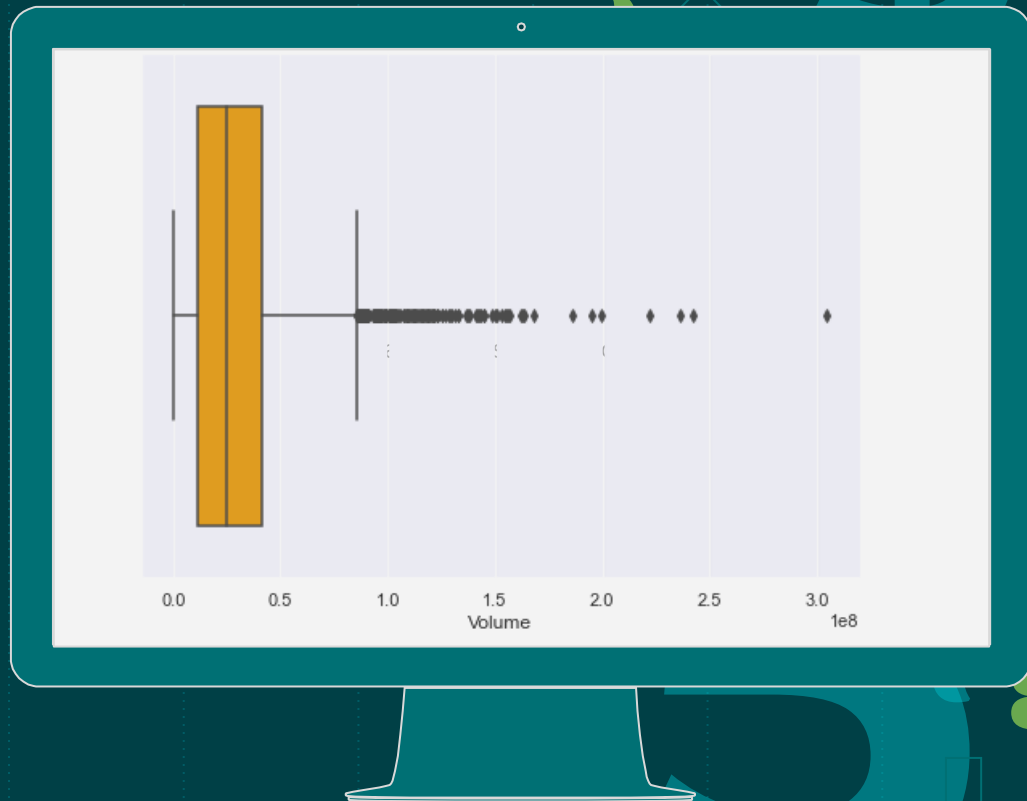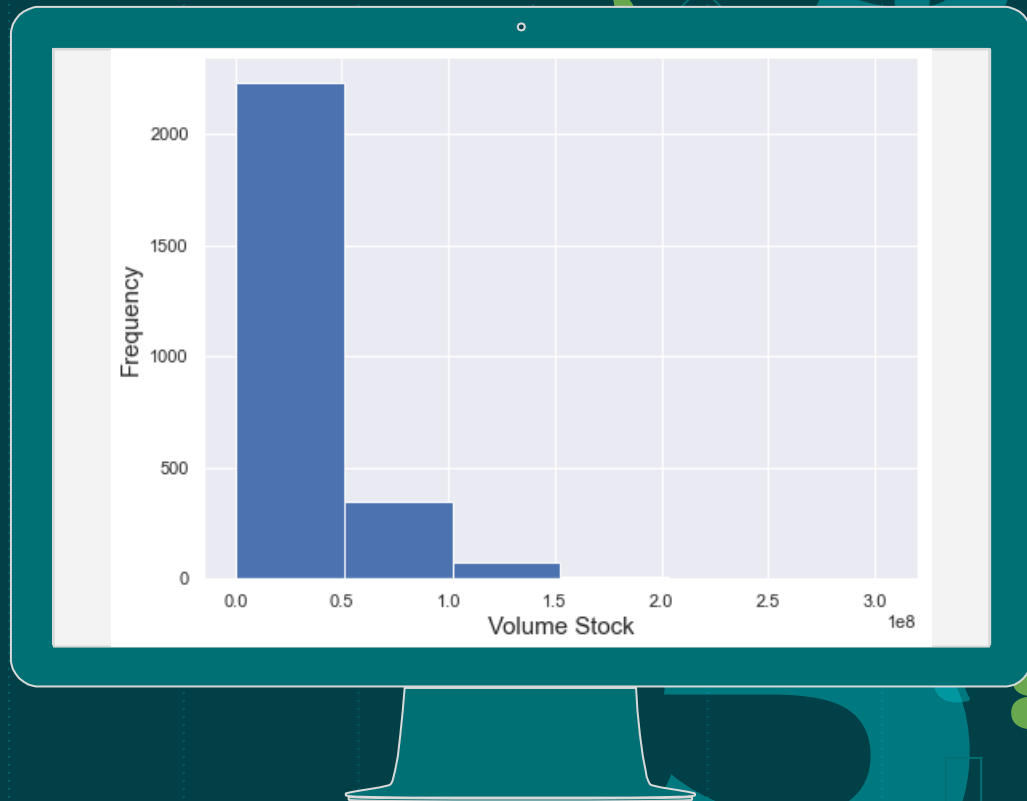
# Descriptive Analysis

─○ **Univariate Analysis**

Showing the targe varible Close Price using the linear graph



Tesla Close Price History

# Descriptive Analysis

—○ **Univariate Analysis**

Volume Frequency

# Descriptive Analysis

○ **Multivariate Analysis**

Varibales Relationships
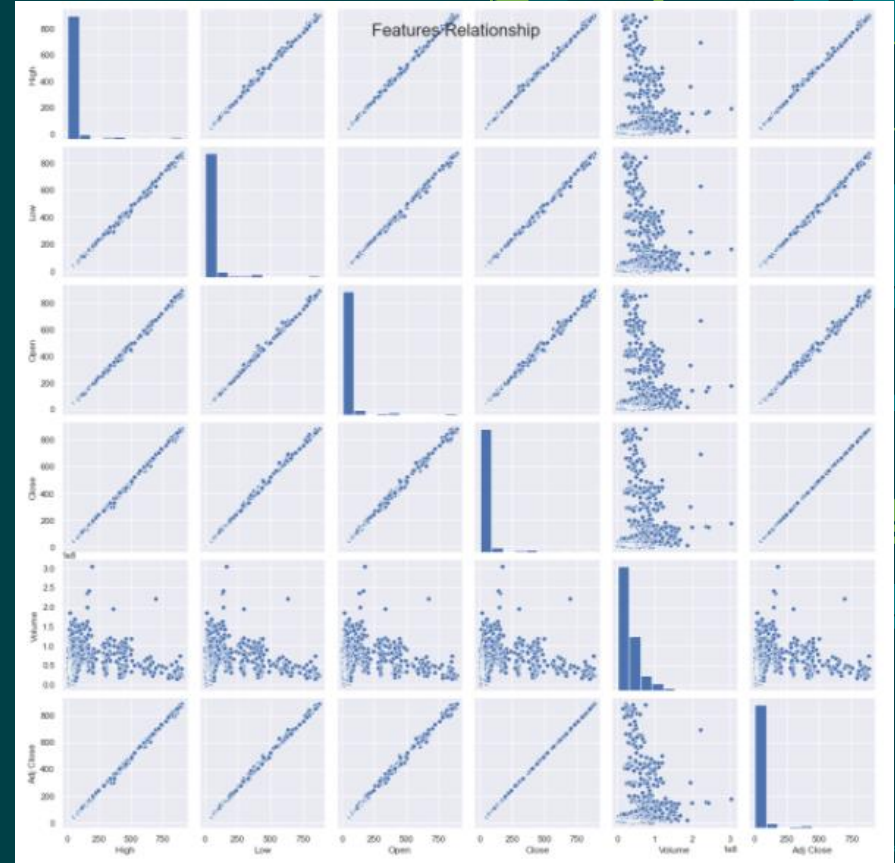
# Descriptive Analysis

─○ **Multivariate Analysis**

Show and explain your web, app or software projects using these gadget templates.

# Descriptive Analysis

─○ **Multivariate Analysis**

Visualizing outliers in Stock Price Varibles
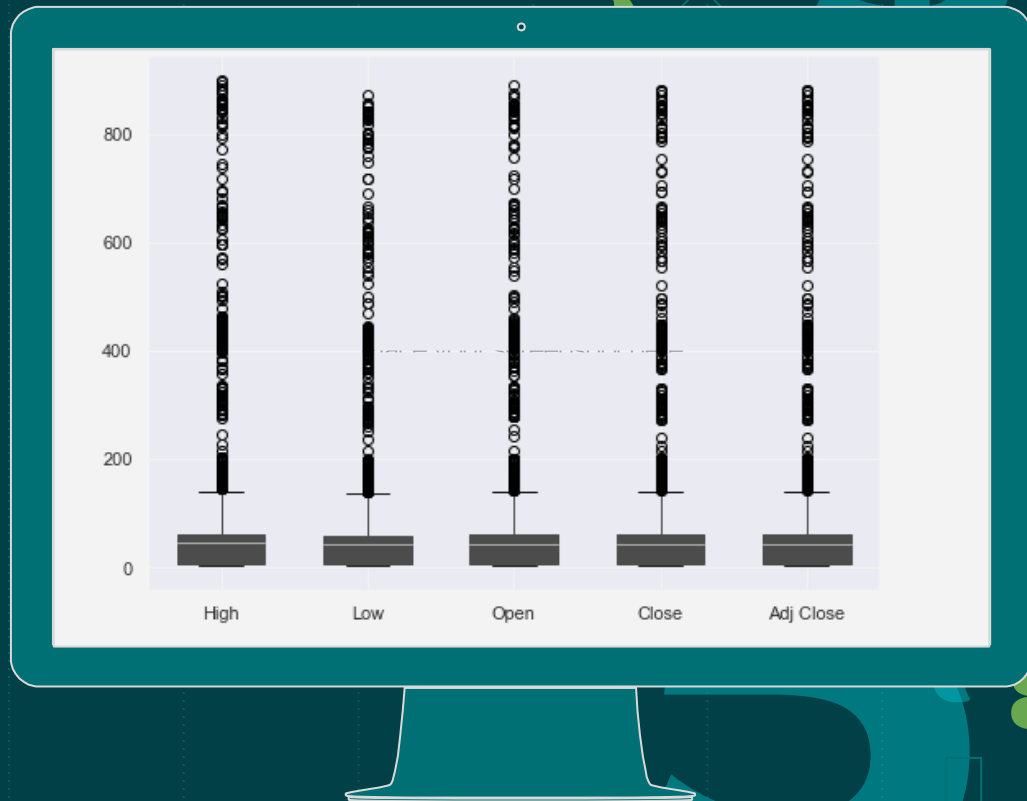
# Assumptions

## Stationary

We trying to know the close price is stationary or not.

```
ADF Statistic: 6.21024818678538
p_value: 1.0

We fail to reject the H0
So, The Close prices is non-Stationary
```

# Assumptions

─○ **Autocorrelation**

Represents the degree of similarity between a given close price
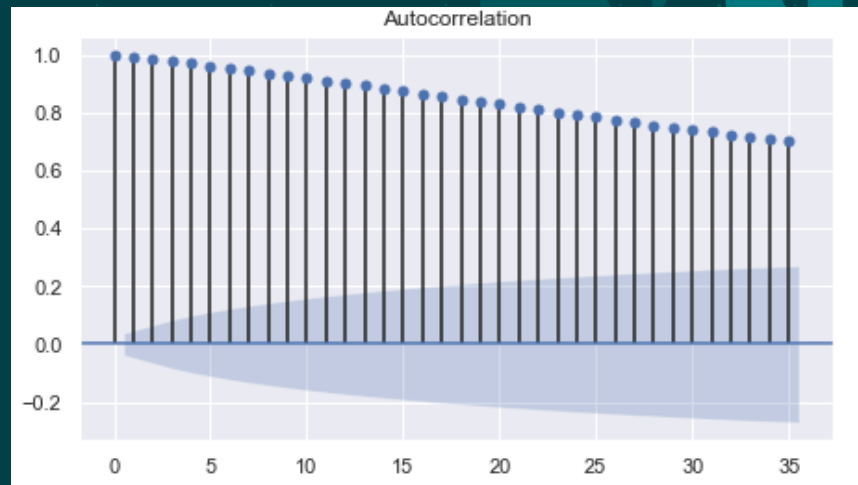
# Close Price Predication Models

- o **Decision Tree Model**

Forecasting the close price and testing

# Close Price Predication Models

- ○ **Linear Regression Model**

Forecast Prediction for close price and validate the testing data
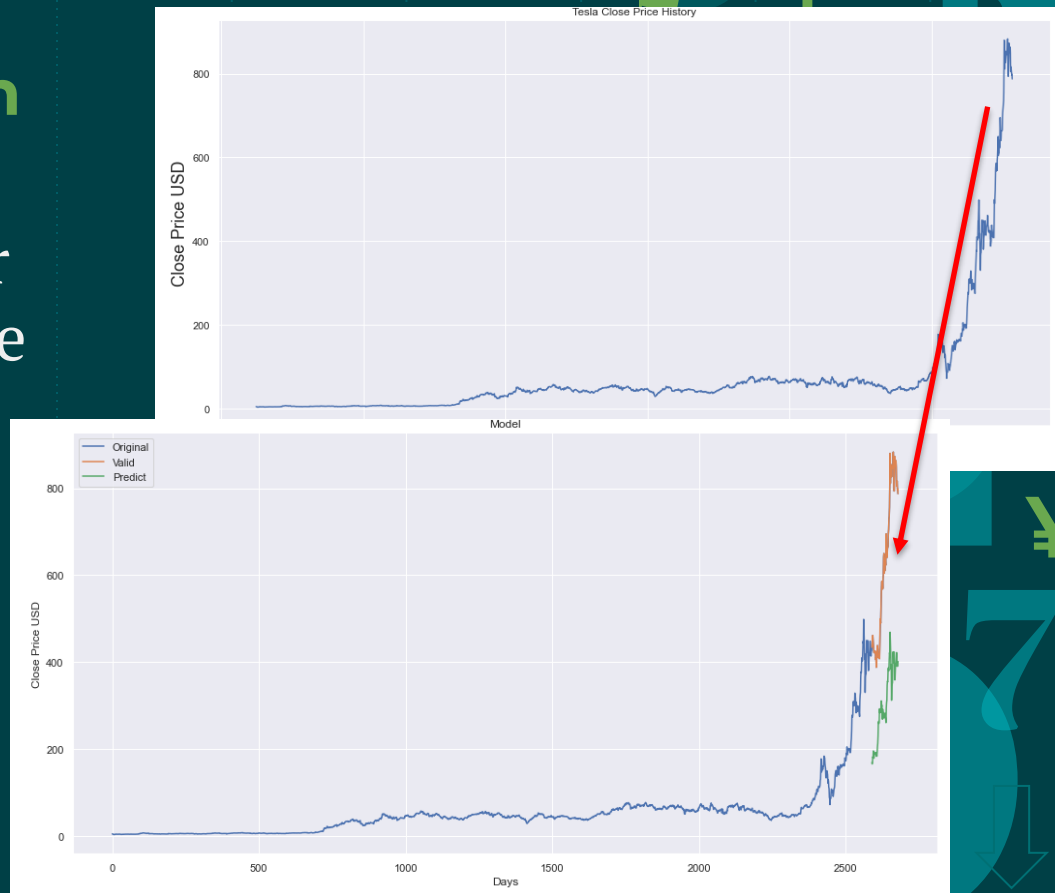
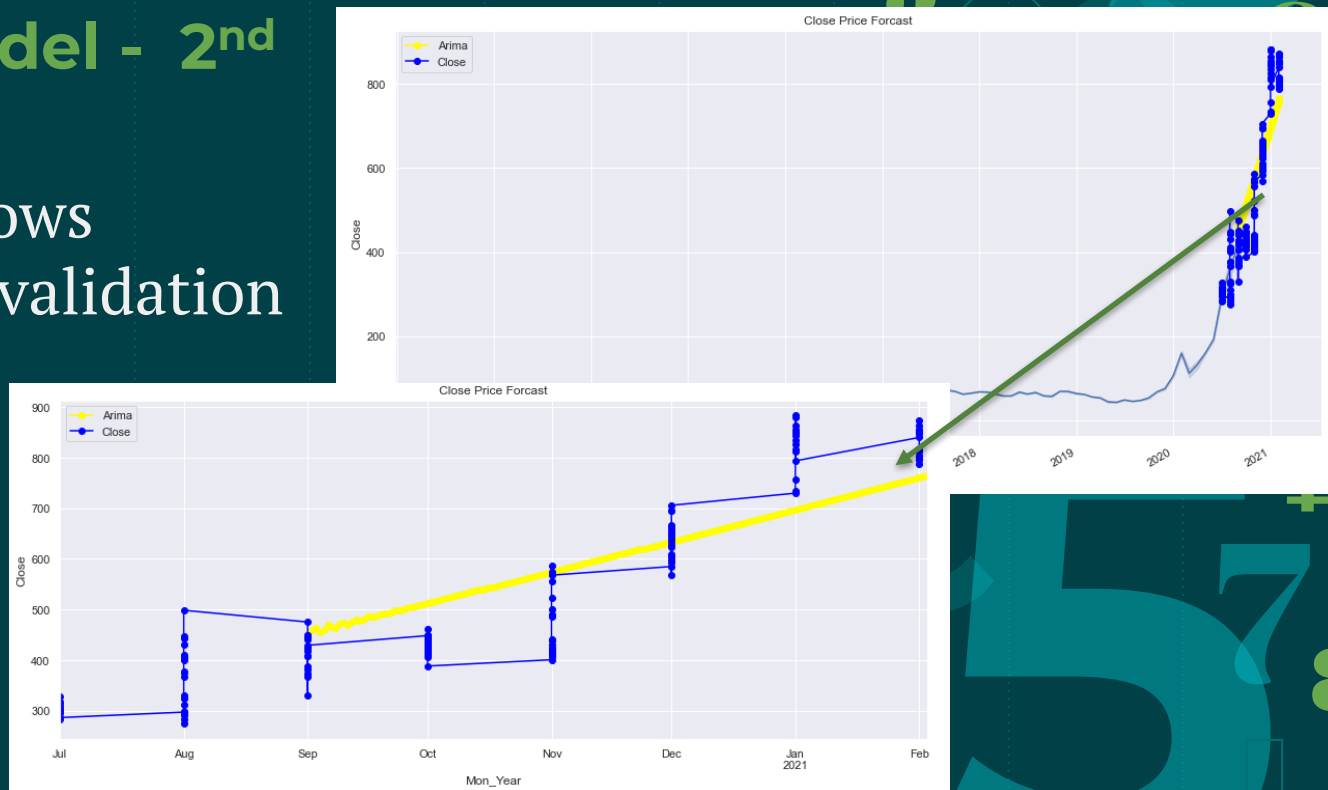# Close Price Predication Models

- **Arima Model – 1st Approach**

The graph shows predicted for validation model

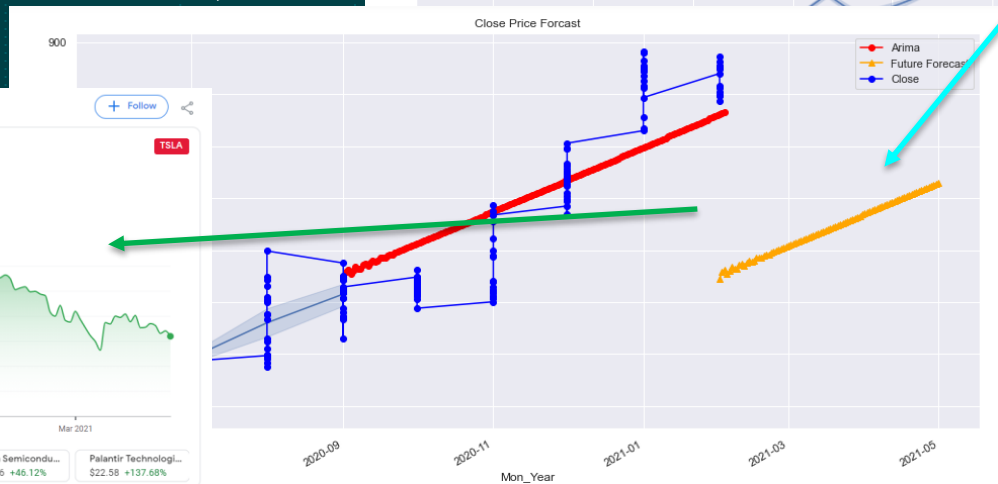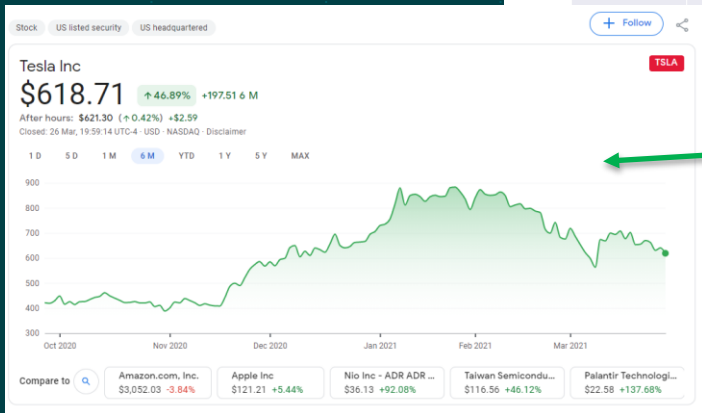# Close Price Predication Models

─○ **Arima Model - 2nd Approach**

The graph shows predicted for validation model
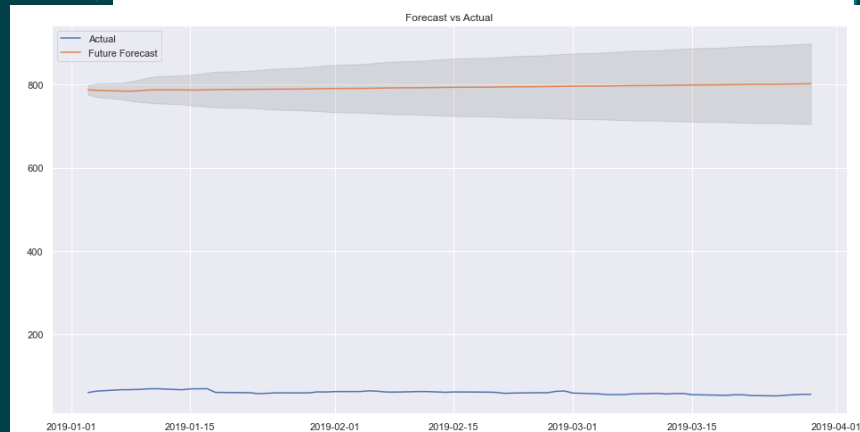
# Close Price Predication Models

—o **Arima Model - 2ⁿᵈ Approach**

Predicted forecast close price for rest of Feb, Mar, 2021

# Conclusions/Findings

- The Close Price are Autocorrelated and Non-Stationary

- All the variables are highly correlated to target variable

- The best model to predict the foresting for Close price is Arima Model

- Model observed that the price is dropped down which is true predicted (as I showed in slide 22), still it is not achieved the but it very close.

# Thanks!

Any questions?